

В.А. Нуриев

ДВОЕТОЧИЕ: ФУНКЦИОНАЛЬНАЯ (А)СИММЕТРИЯ МЕЖДУ РУССКИМ И ФРАНЦУЗСКИМ ЯЗЫКАМИ

Институт научной информации по общественным наукам (НИИОН)
Российской академии наук, Москва, Россия; nurievv@mail.ru

Аннотация. В статье представлено исследование функционального потенциала двоеточия на основе корпусных данных в языковой паре русский-французский. Контрастивный анализ двоеточия выполнен в соответствии со специально разработанной структурной схемой. При этом решены следующие задачи: рассчитана частотность двоеточия в сопоставляемых языках, уточнена его функциональная нагрузка, выявлены зоны функциональной симметрии и асимметрии. Основной корпусный информационный инструмент — база данных, функционирующая на основе параллельных корпусов Национального корпуса русского языка. Она предусматривает возможность формирования необходимых поисковых запросов для обработки пунктуационного компонента параллельных текстов. Корпусно-ориентированные контрастивные исследования пунктуации непосредственно связаны с актуальными задачами информатики. Они необходимы для проектирования ряда современных информационных продуктов на основе искусственного интеллекта, в частности, голосовых ассистентов нового поколения. Также их результаты позволят усовершенствовать существующие технологии машинного перевода. Они помогут настроить и отладить обучение систем машинного перевода так, чтобы учитывались пунктуационные характеристики, не совпадающие в контактирующих при переводе языках. Кроме того, эти результаты будут востребованы при развитии технологий субтитрирования на основе искусственного интеллекта.

Ключевые слова: пунктуация, двоеточие, межъязыковая асимметрия, корпусные информационные инструменты, параллельный корпус, русский, французский

doi: 10.55959/MSU-2074-1588-19-28-3-2

Нуриев Виталий Александрович — доктор филологических наук, ведущий научный сотрудник Института научной информации по общественным наукам (НИИОН) Российской академии наук; nurievv@mail.ru

© Нуриев В.А., 2025



Для цитирования: Нуриев В.А. Двоеточие: функциональная (а)симметрия между русским и французским языками // Вестн. Моск. ун-та. Сер. 19. Лингвистика и межкультурная коммуникация. 2025. Т. 28. № 3. С. 28–40. DOI: 10.55959/MSU-2074-1588-19-28-3-2

Введение

Цель статьи — контрастивное изучение функционального потенциала двоеточия в языковой паре французский-русский. Для ее достижения необходимо было изучить, насколько частотен этот знак в сопоставляемых языках, какую функциональную нагрузку он несет, и выявить зоны функциональной (а)симметрии.

Трехчастная структура, соответствующая трем основным разделам статьи, определена в ходе обработки эмпирических данных с опорой на [Ström Herold, Levin, 2021; Pecorari, Longo, 2021]. По этой структурной схеме были выполнены контрастивный анализ и описание функционального потенциала двоеточия в сопоставляемых языках.

Материал и метод

Основным исследовательским инструментом является база данных, функционирующая на основе параллельных корпусов Национального корпуса русского языка (НКРЯ) и предусматривающая возможность формирования поисковых запросов для обработки пунктуационного наполнения параллельных текстов. Таким образом привлекается эмпирический корпусный материал представительного объема — более 7 млн словоупотреблений в двух параллельных корпусах: французско-русский — 2 459 749, русско-французский — 4 684 900¹. Исходные данные удовлетворяют принципу сопоставимости, то есть речевой материал относительно равномерно распределен по целому ряду параметров: жанровое разнообразие, количественное равновесие и т.д. В каждом переводном направлении предварительно были определены 2 основные категории: художественная и нехудожественная литература. Объем обрабатываемого материала в обеих категориях был приведен в соответствие, чтобы избежать искажения результатов. Всего обработано 21 000 параллельных контекстов. На этом материале изучался функциональный объем двоеточия.

¹ В статье в силу ограниченного объема приводится только анализ русско-французского корпусного материала. Анализ французско-русского корпусного материала будет приведен в дальнейших публикациях.

Задавались следующие поисковые запросы:

- 1) двоеточие в исходном тексте (ИТ); двоеточие в переводном тексте (ПТ);
- 2) двоеточие в ИТ; иной пунктуационный знак в ПТ;
- 3) двоеточие в ИТ; отсутствие какого-либо пунктуационного знака (нулевой эквивалент) в ПТ;
- 4) иной пунктуационный знак в ИТ; двоеточие в ПТ;
- 5) отсутствие какого-либо пунктуационного знака (нулевой эквивалент) в ИТ; двоеточие в ПТ.

Эти запросы определили последующую инвентаризацию данных. Все найденные контексты сгруппированы согласно структуре моделей перевода, которая соотносится с четырьмя типами переводческих операций — сохранение, замена, опущение (может трактоваться как замена исходного знака препинания нулевым эквивалентом), добавление. Моделью перевода при этом считается кортеж вида <языковая единица; {вариант перевода}>, где {вариант перевода} обозначает весь спектр вариантов, приемлемых для перевода конкретной языковой единицы.

Методологические принципы контрастивного исследования пунктуации обозначены в [Нуриев, Игнатова, 2024; Нуриев, Карпов, 2023; Нуриев, Кружков, 2023; Nádvorníková, 2020].

Пунктуационные нормы и узус в сопоставляемых языках

В разделе представлены нормативные и узуальные установки при употреблении двоеточия. Функциональный потенциал знака препинания описывается согласно нормирующим источникам (грамматические справочники и дидактическая литература), также на корпусном материале выделяются узуальные — регулярно воспроизведимые — случаи употребления. Иллюстративный материал приводится в форме параллельных контекстов, чтобы показать межъязыковую (а)симметрию в употреблении изучаемого знака. Ниже приводится фрагмент полученного описания на примере русского двоеточия на фоне французского языка. Для его составления были использованы следующие нормирующие источники: [Валгина, 1979; Валгина, 2004; Правила русской орфографии и пунктуации, 2007; Шапиро, 1955; Шапиро, 1974]. Описание функционального потенциала французского двоеточия на фоне русского языка по представленной схеме не приводится здесь в силу ограниченного объема статьи.

Употребление двоеточия в русском языке, как и в целом пунктуирование, подчинено семантико-грамматическому принципу. Двоеточие выполняет одновременно демаркативную и смыслоразличительную функции, отмечая границу двух частей предложения (Р и Q), между которыми устанавливаются определенные семантические отношения. Сфера его речевого обихода прежде всего затрагивает бессоюзие: на границе предикативных частей бессоюзного сложного предложения оно является сигналом последующего разъяснения с различными смысловыми оттенками и ставится в следующих случаях.

Q содержит пояснение того, о чём говорится в Р (между частями уместна постановка *а именно*).

(1)

Вдруг в глазах моих совершилось невероятное дело: отец внезапно поднял хлыст, которым сбивал пыль с полы своего сюртука, — и послышался резкий удар по этой обнаженной до локтя руке. [И.С. Тургенев. Первая любовь (1860)]

Et il se produisit alors une chose incroyable : mon père leva brusquement sa cravache, avec laquelle il tapait sur les pans poussiéreux de sa veste, et cingla violement le bras de la jeune fille, nu jusqu'au coude. [Tr. M.-R. Hofmann, 1974]

Р содержит обобщающее слово типа *так*, *таков*, *такой*, *одно* и т.д., содержание которого конкретизирует Q.

(2)

Практика была такая: ребенок рождается — его сразу отдают в деревню. [С. Алексиевич. Время секонд хэнд (ч. 1) (2013)]

Et on s'y prenait comme ça : dès qu'un bébé naissait, on le laissait dans un village. [Tr. S. Benech, 2013]

Q указывает причину положения вещей, описываемого в Р (между предикативными частями возможна постановка союза типа *так как*, *потому что*).

(3)

Рядом мальчик с округлившейся щекой: внутри угадывается конфета. [С. Довлатов. Иностраница (1986)]

À côté d'elle un petit garçon à la joue proéminente : à l'intérieur on devine un bonbon. [Tr. J. Michaut-Paterno, 2001]

Q раскрывает содержание Р, сообщая дополнительный факт (предикативные части связывают изъяснительно-пояснительные отношения, подчеркнутые имеющимся в Р глаголом чувственного или рационального восприятия типа *знать*, *узнать*, *думать*, *понимать*, *помнить*, *видеть*, *слышать*, *смотреть*, *чувствовать* и т.д., а между Р и Q возможна постановка союза *что*).

(4)

Ведь он знает: Садко меня не любит. [В. Гроссман. Жизнь и судьба, ч. 2 (1960)]

Tu comprends, il le sait : Sadko ne m'aime pas. [Tr. A. Berelowitch, 1980]

При сходных семантико-синтаксических условиях в Р отсутствует перцептивный глагол, указывающий на результат восприятия и способный непосредственно присоединять изъяснение (типа *увидел, почувствовал, услышал, что...*), однако имеются другие глаголы или глагольные конструкции, предвосхищающие следующий этап восприятия, но синтаксически ограниченные в построении союзных придаточных предложений со значением изъяснения (типа *взглянуть, оглянуться, прислушаться, посмотреть* и т.п.).

(5)

Он взглянул: в правой руке у него отрезанные куски бахромы, носок и лоскутья вырванного кармана. [Ф.М. Достоевский. Преступление и наказание (1866)]

Il regarda : dans sa main droite il tenait les morceaux de la frange, la chaussette et les lambeaux de la poche arrachée. [Tr. É. Guertik, 1947]

Q представляет собой вопрос, задаваемый себе или собеседнику, а в Р имеется слово, указывающее на формальные характеристики вопроса: мысленный или вербально выраженный, в форме просьбы или требования и т.д. Этот тип бессоюзного сложного предложения сближается с прямой речью, отличаясь тем, что Q передает самую суть вопроса, но без характерных для прямой речи атрибутов — обращений, частиц и т.п.

(6)

Я себя спрашиваю: в самом деле, не полечиться ли у кого-нибудь из товарищей? [А.П. Чехов. Скучная история (1889)]

Je m'interroge : ne devrais-je pas effectivement me faire soigner par un de mes collègues ? [Tr. Édouard Parayre, 1960]

(7)

— Да, Афраиний, вот что мне внезапно пришло в голову: не покончил ли он сам с собой? [М. Булгаков. Мастер и Маргарита (ч. 2) (1929–1940)]

— Mais dites-moi, Afranius. Il me vient soudain une idée : n'aurait-il pas lui-même mis fin à ses jours? [Tr. C. Ligny, 1968]

В сложноподчиненном предложении двоеточие ставится в двух следующих случаях (примеры 8–9).

Если Р содержит слова, сигнализирующие о последующем пояснении, и возможна подстановка *а именно*.

(8)

Господь подарил нам жизнь и хочет от нас одного: чтобы мы осознали, для чего он одарил нас этой самой жизнью. [В.Г. Сорокин. Метель (2010)]

Dieu nous fait don de la vie et n'exige de nous qu'une chose : que nous ayons une claire conscience de ce qui motive notre présent. [Tr. A. Coldefy-Fauvard, 2011)]

Если Q представляет собой косвенный вопрос.

(9)

Ибо, вернувшись из Чистополя, Цветаева — это теперь очевидно! — колеблется: стоит ли уезжать из Елабуги? [И. Кудрова. Путь комет. Разоблачённая морока (2007)]

Car Tsvétaïeva, revenue de Tchistopol — cela est aujourd'hui évident! — hésite : est-ce bien la peine de quitter Élabouga? [Tr. H. Hengy, 2015]

В простом предложении двоеточие также реализует разъясняющий семантический потенциал, конкретизируя значение обобщающего слова, помещенного перед сочинительной конструкцией (однородными членами).

(10)

Разумеется, у Черкасовых были друзья из высшего социального круга: Шостакович, Мравинский, Эйзенштейн... [С. Довлатов. Чемодан (1986)]

Bien entendu, les amis des Tcherkassov appartenaient à la crème : Chostakovitch, Mavrinski, Eisenstein... [Tr. Jacques Michaut-Paterno, 2001]

(11)

С одной стороны — деятели политики, государственного аппарата: посол России во Франции Извольский, сенатор Куиба, депутат Стег, управляющий из МИДа Бапст, сам Думер, бывший министр и бывший президент Палаты депутатов и т.д. [www.ifspb.com. Сайт Французского института в Санкт-Петербурге [ABBYY LingvoPRO] (2002–2006)]

D'un côté, les hommes politiques, l'appareil d'État : l'ambassadeur de Russie en France Izvolski, le sénateur Couyba, le député Steeg, un directeur du Ministère des Affaires étrangères russe, Bapst, Doumer lui-même, ex-ministre et ex-président de la Chambre des Députés, etc. [www.ifspb.com. Le site de l'Institut Français à Saint-Pétersbourg (2002–2006)]

Двоеточие ставится, когда за обобщающим словом следуют уточняющие *как например, например, как то, а именно*. Перед ними ставится запятая, а после них — двоеточие.

(12)

Из преступлений одно, именно: кражи гороху, моркови и репы по огородам, — было в большом ходу, да однажды вдруг исчезли два поросенка и курица — происшествие, возмущившее весь околоток и приписанное единогласно проходившему накануне обозу с деревянной посудой на ярмарку. [И.А. Гончаров. Обломов (1848–1859)]

De tous les délits, on n'en connaissait qu'un seul : le vol des pois, des carottes et des navets dans les potagers ; une fois seulement une poule et des porcelets avaient disparu. Cet événement avait indigné tout le district et chacun le rapprocha du passage d'un convoi de charretiers menant à la foire de la vaisselle de bois. [Tr. A. Adamov, 1959]

В предложениях с сочиненными рядами при отсутствии обобщающего слова возможна постановка двоеточия в качестве сигнала последующего перечисления, перечня. Эту сигнальную функцию двоеточие выполняет при оформлении перечней в официальных документах.

(13)

Статья 102 1. К ведению Совета Федерации относятся: а) утверждение изменения границ между субъектами Российской Федерации; б) утверждение указа Президента Российской Федерации о введении военного положения; в) утверждение указа Президента Российской Федерации о введении чрезвычайного положения... [Конституция Российской Федерации (1993)]

Article 102 1. Relèvent de la compétence du Conseil de la Fédération : a) l'approbation de la modification des frontières entre les sujets de la Fédération; b) l'approbation du décret du Président de la Fédération de Russie sur l'introduction de l'état de siège; c) l'approbation du décret du Président de la Fédération de Russie sur l'introduction de l'état d'urgence... [Tr. M. Lesage, 2000]

Двоеточие оформляет конструкции с прямой речью, следуя за словами автора:

(14)

Маруся удивилась: — Я и не подозревала, что вы знаете такие слова. [С. Довлатов. Иностраница (1986)]

Maroussia fut tout étonnée : « Je n'aurais jamais cru que vous connaissez des mots pareils. »

[Tr. J. Michaut-Paterno, 2001]

Двоеточие употребляется в двучленных конструкциях заголовочного типа с именительным (и винительным) темы, являясь своеобразным эмфатизатором-разъяснителем.

(15)

В самолет мне не разрешили ничего взять, только сумочку с документами. Ни вещи, ни мамины пирожки: «Поймите — военная обстановка». [С. Алексиевич. Время секонд хэнд (ч. 1) (2013)]

Dans l'avion, on ne m'a rien laissé emporter, juste mon sac à main avec mes papiers. Rien d'autre, pas même les petits pâtés que maman m'avait donnés. « On est en guerre, vous comprenez. » [Tr. S. Benech, 2013]

В примере 15 представлен случай межъязыковой асимметрии: русское двоеточие в переводе заменено концевой точкой, что приводит к декомпозиции исходной конструкции. Это тем более примечательно, что во французском имеется сходная функция исследуемого знака. Он также может участвовать в двухчленных конструкциях заголовочного типа, отмечая границу частей и являясь эмфатизатором-разъяснителем, который позволяет значительно экономить синтаксические средства выражения вне грамматических конвенций. При этом Р содержит именную группу, где обозначена тема, а Q эту тему развивает, поясняет, комментирует:

(16)

Attentat à Paris : le terroriste avait prévu d'attaquer une cible juive. [Le Monde, 8 Dec. 2023, p. 13]

Пунктуационный узус охватывает два типа воспроизводимых контекстов с двоеточием.

В первом случае двоеточие сигнализирует о противопоставлении, заключенном в Q.

Во втором — об оценочном суждении о Р, представленном в Q.

Ни тот, ни другой случаи не описаны в нормирующей литературе, однако примеры такого употребления зафиксированы уже в текстах XIX в.

(17)

— Ты же от почтальона принял его: грязное такое! [И.А. Гончаров. Обломов (1848–1859)]

— C'est toi qui l'as reçue du facteur : une lettre toute sale. [Tr. L. Jurgenson, 1988]

(18)

Я баб погнал по мужей: бабы те не воротились... [И.А. Гончаров. Обломов (1848–1859)]

J'ai envoyé les bonnes femmes chercher leurs maris : lesdites bonnes femmes ne sont pas revenues... [Tr. L. Jurgenson, 1988]

Согласно типографическим требованиям, в русском языке двоеточие отделяется от последующего слова пробелом, в то время как, например, во французском языке оно отделяется пробелами и от предшествующего, и от последующего слов.

Результаты контрастивного корпусного исследования

Частотность употребления двоеточия в сопоставляемых языках составляет (данные получены на основе запросов в моноязычных корпусах: XX век и современный корпусы Frantext (157 314 064 словоупотреблений) — для французского, основной корпус НКРЯ (374 449 975 словоупотреблений) — для русского):

- 1) русский язык — 2 535 515 двоеточий всего и 68 на 10 000 слов;
- 2) французский язык — 801 732 двоеточий всего и 51 на 10 000 слов.

Таким образом, в русском языке двоеточие продуктивнее, чем во французском, что подтверждается данными параллельных корпусов.

Зона функциональной симметрии в области пунктуирования между сопоставляемыми языками значительна, однако по направлениям перевода она совпадает не полностью:

- при переводе с французского на русский двоеточие сохраняется в соответствующей синтаксической позиции в 70,26% случаев;
- при переводе с русского на французский двоеточие сохраняется в соответствующей синтаксической позиции в 62,67% случаев.

В зоне симметрии не полностью совпадает и распределение функционального веса двоеточия как знака-носителя определенной семантической нагрузки.

Например, и русский, и французский преимущественно используют двоеточие для оформления прямой, несобственно-прямой и косвенной речи во всем ее многообразии (включая цитаты, правила, примеры, инструкции, объявления, вывески и прочую справочную информацию). Однако в этой функциональной области русский язык сохраняет двоеточие при переводе с французского в 59,22%, а французский при переводе с русского — в 64,89%.

В сфере бессоюзия в переведенном русско-французском направлении симметричная постановка наблюдается для разъяснительного двоеточия — в 19,13% случаев, для причинного — в 6,49% случаев, для аргументативного (вводящего дополнительный аргумент в пользу Р) — в 2,53% случаев, для следственного — в 1,69% случаев. Во французско-русском направлении перевода цифры несколько разнятся: симметричные синтаксические условия наблюдаются в 27% случаев для постановки разъяснительного двоеточия, в 4,25% случаев — для причинного, в 3,65% случаев — для аргументативного, в 0,47% случаев — для следственного двоеточия.

Любопытно, что в обоих переводных направлениях зона симметрии покрывает малое число случаев, где двоеточие является входным знаком сочиненного ряда: 3,94% (русско-французское) и 5,61% (французско-русское), причем нормирующая литература всегда делает акцент на этой функции двоеточия.

Зона пунктуационной асимметрии определялась посредством четырех из пяти поисковых запросов, приведенных выше в разделе «Материал и метод» (запросы 2–5). Так, было выявлено, что при переводе с французского на русский двоеточие заменяется иным пунктуационным знаком в 26,51% случаев, а опускается в 3,22% случаев, наличие в ИТ какого-либо пунктуационного знака служит стимулом для реализации двоеточия в ПТ в 15,91% случаев, постановка двоеточия в ПТ при отсутствии какого-либо пунктуационного знака-стимула в ИТ происходит в 10,34%. В обратном направлении двоеточие заменяется иным пунктуационным знаком в 33,29% случаев, а опускается в 10,43% случаев, наличие в ИТ какого-либо пунктуационного знака служит стимулом для реализации двоеточия в ПТ в 22,84% случаев, постановка двоеточия в ПТ при отсутствии какого-либо пунктуационного знака-стимула в ИТ происходит в 3,5%.

Показательно, что при переводе с французского на русский число двоеточий возрастает с 55 в оригинале до 71 в переводе (в расчете на 10 000 слов), а в обратном направлении убывает — с 97 до 58 (в расчете на 10 000 слов).

Заключение

Корпусно-ориентированное контрастивное изучение двоеточия позволяет сделать ряд лингвистически интересных выводов. Так, можно утверждать, что двоеточие в русском продуктивнее, чем во французском, и что, будучи медиальным (то есть срединным, тяготеющим к постановке внутри предложения — на границе составных частей) знаком, оно при переводе может заменяться и концевыми знаками препинания при статистически регулярной воспроизведимости такого трансформационного сдвига. Также в указанной языковой паре двоеточие при переводе является хоть и не абсолютным, но существенным пунктуационным прескриптором. В этом смысле дальнейшие перспективы исследования связаны с выявлением и описанием устойчивых переводных соответствий двоеточия, особенностей передачи его функциональной значимости при переводе и основных трансформационных синтаксических сдвигов, сопряженных с этим. Кроме того, подобное изучение пун-

ктуации помогает выявить ряд ограничений. Одно из них, существенно влияющее на валидность данных, — сбалансированность литературной продукции в составе корпуса по количественным (протяженность текстового источника) и качественным характеристикам (жанровое и стилистическое разнообразие текстовых источников). Статистическая обработка данных требует тонкой нюансировки: при подсчете абсолютных числовых величин необходимо, например, учитывать мультивариантную структуру корпусного ресурса, в котором один текст может быть представлен несколькими переводами. Также продуктивность семантически насыщенных знаков препинания может в конкретном произведении зависеть от сюжетной развертки, что, очевидным образом, следует учитывать при анализе корпусных данных.

СПИСОК ЛИТЕРАТУРЫ

1. Валгина Н.С. Русская пунктуация: принципы и назначение. М.: Просвещение, 1979.
2. Валгина Н.С. Актуальные проблемы современной русской пунктуации: Учеб. пособие. М.: Высш. шк., 2004.
3. Нуриев В.А., Игнатова С.Д. Надкорпусная база данных как инструмент изучения пунктуации // Вестник Моск. ун-та. Сер. 19. Лингвистика и межкультурная коммуникация. 2024. Т. 27. № 4. С. 147–158.
4. Нуриев В.А., Карпов В.И. Методология корпусно-ориентированного исследования в области контрастивной пунктуации // Информатика и ее применение. 2023. Т. 17. № 2. С. 90–95.
5. Нуриев В.А., Кружков М.Г. Корпусные данные при контрастивном изучении пунктуации. Системы и средства информатики. 2023. Т. 33. Вып. 1. С. 14–23.
6. Правила русской орфографии и пунктуации. Полный академический справочник / ред. В.В. Лопатин. М.: Эксмо, 2007.
7. Шапиро А.Б. Основы русской пунктуации. М.: Изд-во Академии наук СССР, 1955.
8. Шапиро А.Б. Современный русский язык. Пунктуация. М.: Просвещение, 1974.
9. Nádvorníková O. The use of English, Czech and French punctuation marks in reference, parallel and comparable web corpora: a question of methodology // Linguistica Pragensia. 2020. Vol. 30. Iss. 2. Pp. 30–50.
10. Pecorari F., Longo F. The Ellipsis and the Dash in Italian and English: A Contrastive Perspective // Vergleichende Interpunktions — Comparative Punctuation / Eds. P. Rössler, P. Besl, A. Saller. Berlin, Boston: De Gruyter, 2021. Pp. 289–314.
11. Ström Herold J., Levin M. The Colon in English, German and Swedish: A Contrastive Corpus-Based Study // Vergleichende Interpunktions — Comparative Punctuation / Eds. P. Rössler, P. Besl, A. Saller. Berlin, Boston: De Gruyter, 2021. Pp. 237–261.

Vitaly A. Nuriev

COLON: FUNCTIONAL (A)SYMMETRY BETWEEN RUSSIAN AND FRENCH

Institute of Scientific Information for Social Sciences of the Russian Academy of Sciences (INION RAN), Moscow, Russia; nurievv@mail.ru

Abstract. The article presents the corpus-assisted research of the functional potential of the colon in the Russian-French language pair. For the data description, a three-part structure was determined. Following that structural model, a contrastive analysis of the colon was performed in order to uncover and then describe its functional potential in the studied language pair, which allowed to solve the following tasks: 1) to calculate the frequency of the colon in the compared languages; 2) to clarify its functional potential there; 3) to identify the zones of functional symmetry and asymmetry. The main information tool is a supra-corpora database that stores parallel texts from the Russian National Corpus. It allows for generating the search queries to process the punctuation component in parallel texts. Corpus-based contrastive studies of punctuation are directly related to current challenges in computer science. They are essential for designing a range of modern AI-based information products, particularly next-generation voice assistants. Additionally, their findings will contribute to improving machine translation technologies. They will aid in calibrating and refining the training of machine translation systems to account for punctuation differences between source and target languages. Moreover, these results will be of use in advancing AI-driven subtitling technologies.

Keywords: punctuation; colon; asymmetry between languages; corpus-based tools; parallel corpus; Russian; French

For citation: Nuriev V.A. (2025) Colon: functional (a)symmetry between Russian and French. *Lomonosov Linguistics and Intercultural Communication Journal*, no. 3(28), pp. 28–40. DOI: 10.55959/MSU-2074-1588-19-28-3-2 (In Russ.)

About the author: Vitaly A. Nuriev — Dr. Habil. in Philology, Lead Researcher, Institute of Scientific Information for Social Sciences of the Russian Academy of Sciences (INION RAN); nurievv@mail.ru.

REFERENCES

1. Valgina N.S. 1979. *Russkaya punktuatsiya: printsipy i naznachenie* [Russian punctuation: principles and purpose]. Moscow, Prosveshchenie Publs. (In Russ.)
2. Valgina N.S. 2004. *Aktual'nye problemy sovremennoy russkoy punktuatsii: Uchebnoe posobie* [Current issues in modern Russian punctuation: A textbook]. Moscow, Vysshaya Shkola. (In Russ.)

3. Nuriev V.A., Ignatova S.D. 2024. Nadkorpusnaya baza dannykh kak instrument izucheniya punktuatsii [Supracorpora database as a tool for studying punctuation]. *Lomonosov Linguistics and Intercultural Communication Journal*, 2024, vol. 27, no. 4, pp. 147–158. (In Russ.)
4. Nuriev V.A., Karpov V.I. 2023. Metodologiya korpusno-orientirovannogo issledovaniya v oblasti kontrastivnoi punktuatsii [The methodology of the corpus-based studies in the field of contrastive punctuation]. *Informatika i ee Primeneniya*, vol. 17, no. 2, pp. 90–95. (In Russ.)
5. Nuriev V.A., Kruzhkov M.G. 2023. Korpusnye dannyye pri kontrastivnom izuchenii punktuatsii [The parallel corpora perspective on studying contrastive punctuation]. *Sistemy i Sredstva Informatiki*, vol. 33, no. 1, pp. 14–23. (In Russ.)
6. Lopatin V.V. (ed.). 2007. *Pravila russkoy ortografi i punktuatsii. Polny akademicheskiy spravochnik* [Russian Orthography and Punctuation Rules: A Comprehensive Academic Guide]. Moscow, Eksmo. (In Russ.)
7. Shapiro A.B. 1955. *Osnovy russkoy punktuatsii* [Fundamentals of Russian Punctuation]. Moscow, Publishing House of the USSR Academy of Sciences. (In Russ.)
8. Shapiro A.B. 1974. *Sovremennyi russkiy yazyk. Punktuatsiya* [The modern Russian language. Punctuation]. Moscow, Prosveshchenie Publs. (In Russ.)
9. Nádvorníková O. 2020. The use of English, Czech and French punctuation marks in reference, parallel and comparable web corpora: a question of methodology. *Linguistica Pragensia*, vol. 30, no. 2, pp. 30–50.
10. Pecorari F., Longo F. 2021. The Ellipsis and the Dash in Italian and English: A Contrastive Perspective. In P. Rössler, P. Besl, and A. Saller (eds.). *Vergleichende Interpunktions — Comparative Punctuation*. Berlin, Germany; Boston, MA, USA, De Gruyter, pp. 289–314.
11. Ström Herold J., Levin M. 2021. The Colon in English, German and Swedish: A Contrastive Corpus-Based Study. In P. Rössler, P. Besl, and A. Saller (eds.). *Vergleichende Interpunktions — Comparative Punctuation*. Berlin, Germany; Boston, MA, USA, De Gruyter, pp. 237–261.

Статья поступила в редакцию 14.05.2025;
одобрена после рецензирования 14.06.2025;
принята к публикации 27.06.2025.

The article was submitted 14.05.2025;
approved after reviewing 14.06.2025;
accepted for publication 27.06.2025